DOCUMENT RESUME

ED 090 695 EA 006 147

TITLE Program Analysis and Design Support for the Office of Long Range Planning, New York State Education Department.

INSTITUTION New York State Education Dept., Albany. Office of Long Range Planning.; Riverside Research Inst., New York, N.Y.

REPORT NO RRIP-851-120-281
PUB DATE Oct 72
NOTE 96p.

EDRS PRICE MF-$0.75 HC-$4.20 PLUS POSTAGE
DESCRIPTORS Educational Accountability; *Educational Finance; Educational Research; Measurement; *Program Design; *Program Evaluation; *Program Planning; *Reading Programs; Research Proposals; *State Aid
IDENTIFIERS *New York State

ABSTRACT

This report was written in response to requests from the Office of Long Range Planning (OLRP) of the New York State Education Department (SED). The work covers two different areas: the continuation of program analyses and the continuation of program designs addressing problems of educational accountability. As a result of SED resource allocation decisions, OLRP decided to concentrate on an investigation of the effectiveness of alternative methods for awarding categorical funds to school districts and the development of effectiveness measures for reading programs. Because these two tasks relate to different parts of OLRP's program and are different from each other, they are discussed separately. A work plan for each task is proposed in a separate section. (Author/JF)

ED 090695

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRO
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRE
SENT OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

RIVERSIDE RESEARCH INSTITUTE

RIR

80 West End Avenue / New York, New York 10023 / (212) 873-4000

October 1972

PROGRAM ANALYSIS AND DESIGN SUPPORT FOR THE OFFICE OF LONG RANGE PLANNING, NEW YORK STATE EDUCATION DEPARTMENT

Submitted to

Mr. G. L. Freeborne

Assistant Commissioner for Long Range Planning

New York State Education Department

EA 006 147

RRIP/851-120/281

Preface

This proposal is being written in response to requests from the Office of Long Range Planning (OLRP) of the New York State Education Department (SED). The proposed work covers two different areas: the continuation of program analyses and the continuation of program designs addressing problems of educational accountability.

As a result of SED resource allocation decisions, OLRP has decided to concentrate on two work efforts: an investigation of the effectiveness of alternative methods for awarding categorical funds to school districts; and the development of effectiveness measures for reading programs. Because these two tasks relate to different parts of OLRP's program and are very different from each other, they are discussed separately in Sections I and II. In Section III, a work plan for each task is proposed. Section IV describes Riverside Research Institute (RRI), its capabilities, and its key personnel.

Table of Contents

		Page
I.	A Program Analysis of Categorical Funding	1
	A. Problem Statement	1
	1. Situation	1
	2. Project Goal	4
	B. Approach	5
	1. Overview	5
	2. Phase I--Program Analysis	6
	3. Phase II--Program Design	15
	4. Summary	16
II.	Program Design Research: Effectiveness Measures in Reading and for School Operations	18
	A. Background	18
	B. Proposed Effectiveness Measures	25
	1. Reading Effectiveness Measure(s)	25
	2. Orientation Towards Learning: A Measure of the Effectiveness of School Operations	39
	3. Valuing Learning and School: A Measure of the Effectiveness of School Operations	43

Table of Contents (Cont'd.)

	Page
III. Work Plan	49
A. A Program Analysis of Categorical Funding	49
1. Project Goal	49
2. Work Tasks	49
3. Management Plan	51
B. Program Design Research: Effectiveness Measures in Reading and for School Operations	52
1. Project Goal and Objectives	52
2. Work Tasks	52
3. Management Plan	56
Riverside Research Institute Social Systems Division: Key Personnel	58

I. A Program Analysis of Categorical Funding

A. Problem Statement

1. Situation

Categorical grant programs increasingly have become favored devices by which legislative bodies respond to demands that specific problems be addressed. Both the number of categorical funding programs and the dollar amounts involved are growing rapidly. The number of Federal categorical funding programs to State and local governments has increased from 23 in 1945 to an estimated 400 in 1966. The Federal dollar amount rose from $3.3 billion in 1956 to $20.2 billion in 1969, and is expected to rise to $30 billion by 1975. The figure for 1969 represents almost 22% of domestic Federal expenditures. State governments also make heavy use of categorical funding programs through grants to localities--a total of $23.8 billion in 1969.

The expansion of categorical funding programs has been nowhere more evident than in the education of the disadvantaged. Two nearly identical programs, one State, under the Urban Education Act, and one Federal, under Title I of the Elementary and Secondary Education Act, together funded 262.2 million dollars in 1971-72 for projects to benefit disadvantaged students in the local districts of New

York State. Still other programs have the same general goals and procedures.*

The fiscal consequences of these programs have been mixed at best. That the funds have been expended is clear, but whether these funds have been used as supplements to rather than substitutes for local revenues, or whether they have been concentrated on the target groups rather than being diffused to the general student population, are not at all clear.

Measurable results in terms of students' growth are even less encouraging. A recent study** by the American Institutes for Research in the Behavioral Sciences (AIR) is typical of numerous studies of the impact of categorically funded compensatory education programs. This analysis examined 1,200 evaluations from various sources where it had been reported that a project had produced cognitive benefits that were statistically and educationally significant. AIR evaluated these projects and their evaluation reports against the criteria that, to be labeled "successful," a project would have to produce achievement gains for disadvantaged children

---

* E.g., Headstart, Follow Through, Bilingual Education, Dropout Prevention, Adult Basic Education, Upward Bound, Education of the Handicapped.

** Reported in Report on Education Research, Education News Service, October 17, 1971, pp. 6-8.

which were greater than those of their more advantaged counterparts. Further, this rate of gain would have to be maintained until the disadvantaged children actually caught up. The successful project also had to include a representative sample of not less than thirty children, and achievement gains had to be measured by some reliable testing instrument. After applying these criteria, only 326 of the original sample of 1200 survived. Analysis of the remaining projects showed 3.1 percent of these 326 were "successful" in terms of the criterion described above.

The AIR study provides a bleak picture of the projects. Similar analyses have arrived at similar conclusions.

The problem with these studies is that they provide few clues as to "why" a few projects succeed while most fail. That some succeed suggests that compensatory education is not impossible; that most fail suggests that something fundamental is wrong. What is wrong, and why it is wrong, are not addressed. As a consequence, these studies supply little insight into what can be done at any level to increase the effectiveness of categorically funded compensatory education. Moreover, these "evaluation" studies encourage simplistic (but politically popular) conclusions that the problems of the disadvantaged cannot be ameliorated through educational programs.

What is needed is a program analysis which systematically considers alternatives for achieving specific goals in compensatory education. This analysis should include alternative means of conducting categorical funding programs, alternative supplements to categorical funding, and alternatives to categorical funding itself. In New York State, the program analysis should concentrate on ESEA Title I and the Urban Education Act because, together, they account for many times the number of projects or the amount of funds involved in all other categorically funded compensatory education programs combined. For the New York State Education Department (SED), the program analysis should answer the following question: What should SED do to improve the effectiveness of Title I and Urban Education projects--and how should it be done? The proposed study is designed to answer this question, to lead to well-reasoned recommendations for future SED programs and to plans to implement the recommended programs.

2. Project Goal

To support the objectives of Title I and Urban Education, SED does not act directly on disadvantaged students. Instead SED works through local education agencies (LEAs) which plan, manage, and evaluate the funded projects. Other things being equal, the effectiveness of SED's programs depends on the quality of the LEAs' performance. Hence, SED's

programs are undertaken for both immediate and ultimate purposes. The immediate purpuse is to improve the quality of project planning, management, and evaluation, functions currently performed by LEAs. The ultimate purpose is to benefit the educationally disadvantaged. While the ultimate purpose obviously takes priority, fulfillment of the immediate purpose is a necessary intermediate step. The proposed project reflects this formulation in adopting as the goal for the project to determine how SED can improve planning, management, and evaluation of projects funded under Title I and Urban Education. The achievement of this goal will constitute a first step toward improving the education of the disadvantaged, and will provide a framework for recommendations concerning categorical funding as a means for achieving specific educational goals.

B. Approach

1. Overview

The proposed project pursues the goal of determining how SED can improve project planning, management, and evaluation. The project will be performed in two phases:

Phase I - Program Analysis

Phase II - Program Design

Phase I will recommend what SED should do to improve planning, management and evaluation of projects funded under the Title I and Urban Education programs. Phase II will

recommend how SED should carry out the recommendations of Phase I. The two phases will be sequential, but the approach within each will be iterative and as empirical as possible given extant data.

2. Phase I--Program Analysis

Four tasks will comprise Phase I. Task 1 will consist of a description and evaluation of SED's present operations in carrying out the Urban Education and Title I programs. Descriptive data will be drawn from an examination of extant documents and from interviews with SED personnel. This information will be reduced to functional descriptions (flowcharts and prose) of the present operations of the two programs.

Two sorts of evaluation of present Title I and Urban Education programs will be conducted. First will be a process evaluation which will consist of a critique of the functional descriptions of the present SED programs against contemporary principles of planning, management, and evaluation. This evaluation may suggest near-term improvements in SED's processes.

The second evaluation will assess the extent to which SED's present Title I and Urban Education programs are effective in bringing about the immediate objective of encouraging high quality planning, management, and evaluation of LEA projects. Data for this evaluation will be drawn from

a systematic examination of a sample of the project allocations and evaluations submitted to SED by LEAs under the two programs. This second evaluation will provide a base-line for comparisons of alternatives in Tasks 2 and 3.

Task 2 will estimate the relative effectiveness of alternative incremental changes in SED's present programs, leading to a trade-off analysis among these options. The incremental changes that will be examined will consist of alternative combinations of critical design elements in the present programs, for example: the rigor of the requirements for project documentation (plans and evaluations) by LEAs, and whether funds are allocated by formula or by project competition; whether LEAs are expected to assume funding without SED support at some future date.

The first approximation to these estimates will be drawn from a comparison of the effects on LEA performance on several of SED's present programs which differ with respect to these critical design elements. Data for this comparison will be drawn from a systematic examination of a sample of LEA applications and evaluations much like the examination in Task 1, but including categorically funded, compensatory programs other than Title I and Urban Education.* For

---

* This will also provide, as by-products, data for summary evaluations of the selected programs. Should these evaluations reveal that any of the comparison programs have been

example, ESEA Title I and Urban Education differ principally in the rigor of documentation requirements; Bilingual Education is a project grant but Title I and Urban Education are formula grants; and so on.

The first approximation will be refined by constructing and using a crude description of the cause-effect relationships operative in project planning, management, and evaluation in LEAs. Correlations among LEA performance indicators and between LEA performance and such factors as district size, wealth, homogeneity, etc., will contribute to this description. Further refinement will be provided by a series of on-site observations and interviews in a small number of LEAs selected as "extreme cases" of high and low quality planning and evaluations as determined in the examination of LEA documents in the first approximation, described above. The field work will help identify important differences between LEAs having high and low quality performance. The resultant description, actually a very crude model will permit informed judgments as to why some projects fail while others succeed

---

particularly effective in eliciting the desired performance by LEAs, detailed analyses of these effective programs could by undertaken within the proposed work effort (if mutually agreed to by SED and RRI) or as a separately supported effort. Such additional analyses, which could include process evaluation, would be directed toward determining the reasons for the superior performance. The results of the additional analyses could be folded into the recommendations concerning Title I and Urban Education, as well as into recommendations concerning other categorical funding programs.

and what effects on LEA performance can be expected from incremental changes in SED's programs. Specifically, the description will guide the decisions as to whether the results of earlier comparisons can be extrapolated, or whether (and if so, how) they should be modified in making the predictions for a trade-off analysis among the incremental alternatives.

In Task 3, the search for alternatives will be broadened to include non-incremental alternatives and a second trade-off study will be conducted. A preliminary set of non-incremental alternatives is presented below. The development of the description of LEA performance will augment and refine this list. This preliminary set of alternative SED programs, however, excludes "strawman" design concepts as wasteful of resources both in empirical analyses and in subsequent trade-off analyses.

(a) Upgrading the quality of staff performance within the LEAs might be indicated if the above analyses identified a lack of LEA personnel capability as a major source of ineffectiveness. Three particular capabilities are distinguishable: project planning, project execution and project evaluation. Either of the two sub-alternatives below might address any of these functions or any combination of them.

(a.1) Training of LEA personnel might hold promise if it were found that, although LEAs do not presently have the needed capabilities, they contain people (e.g., the present "coordinators") or can a-tract people who can perform these functions well. In this case, SED might assist in the selection of LEA personnel and might conduct LEA personnel training programs at summer workshops, on the job, at branches of the State University, etc.

(a.2) Attaching SED personnel to LEAs might be indicated if LEAs appeared to lack (and to be unable to attract) persons who could conduct project planning, execution, or evaluation, but had staff members who seemed to be receptive to working with SED personnel skilled in these functions. SED personnel could work solely within one LEA, or could serve several. They could interpret their roles as agents of SED providing consultative support to LEAs, or they could interpret their roles as members of the LEA staff.

(b) Increasing the accountability of the LEA Boards and Superintendents might be a reasonable alternative if the data suggested that skilled LEA personnel either were or could be made available, but that other constraints were impeding their performance and that increased visibility would improve their effectiveness. The two sub-alternatives differ in that the first increases accountability to the Regents, while the second increases accountability to the community.

(b.1) Increasing accountability to the Board of Regents as the State authority responsible for the education of the citizens of New York State, might be accomplished by increasing the rigor of the review processes conducted by SED. Such a step might prove promising if the analyses in T sk 2 suggested that projects funded under Urban Education were more effective than those funded under ESEA Title I. There are many ways of increasing the rigor of review processes and further design work would be required to formulate concrete recommendations. Changes in resource allocation procedures might or might not be involved (see alternative (c), below).

(b.2) Increasing accountability of LEAs to their communities might improve effectiveness if it appeared that Boards of Education, Superintendents, and LEA staff members were highly motivated by community sentiment. SED's activities under this sub-alternative might take the form of formal audits of LEA proposals and evaluations. These audits would appraise the validity of the data, analyses, and interpretations of results contained in LEA evaluations, and would include a written report to the LEA Board. This report would affirm that the LEA staff's procedures were valid, or, if they were not, would state why they were not, and provide appropriate, illustrative commentary. The Board and Superintendent would be required to make public the LEA's own

documents, SED's audit report, and, if they desire, their own comments, within a prescribed time period.

(c) Changing the basis for resource allocation in the LEAs might be appropriate if the earlier analyses indicated that LEAs lack the needed personnel capabilities; that the political constraints associated with pre-allocation of funds according to formulas constituted a major limitation on SED's ability to improve LEA performance; or that LEAs were using ESEA Title I and Urban Education funds to meet basic operating costs. This alternative leads to attempts to modify the resource allocation processes in the LEAs by changing the basis for granting funds to LEAs and thereby influencing the LEAs' allocation of these funds to projects.

(c.1) Constraining the LEAs to selecting among proven programs might be effective if many LEAs appear to be using Title I and Urban Ed funds for projects that are highly comparable with respect to the problems they address, the objectives they pursue, and the design concepts they employ. Under this sub-alternative, SED would develop and test "canned" projects for use in LEAs. LEA staff would select among proven projects, and the funds specifically allocated to the LEAs would be just sufficient to implement and operate the chosen projects. No funds would be available for projects other than those developed by SED. SED might

contract with or collaborate with publishers, LEA personnel, and others in developing and testing "canned" projects.

(c.2) Mixing block and project grants to LEAs might be effective if the LEAs' problems and programming needs were found to be diverse, and if LEAs presently have, or can attract and train personnel possessing programming capabilities. Under this sub-alternative, unconstrained or partially constrained block grants would be awarded to LEAs to bring each LEA budget up to the minimum needed to offset basic operating costs. In addition to these block grants, sizeable project grants would be awarded to those LEAs able to demonstrate competence in project planning, execution, and evaluation. The award of project grants could be made through rigorous review processes conducted by SED personnel or by special review panels convened by SED.

(d) Combinations of the foregoing alternatives and sub-alternatives supply a large number of additional options. A few examples follow:

(d.1) The increasing demands on LEA Boards and Superintendents generated by making them more accountable (alternative (b)) could be met in part by upgrading the quality of LEA staff performance within the LEAs (alternative (a)).

(d.2) Upgrading the skills of persons responsible for project execution in the LEAs (alternative (a)) could

be a critical element in making the SED-generated projects (sub-alternative (a.3)) effective at the LEA level.

(d.3) An important condition for implementing sub-alternative (d.2) might be to increase the responsibility of LEA Boards and Superintendents (alternative (b)). This combination would permit appraisal of the performance of LEA staff by comparative analysis, because the projects of proven effectiveness would come to the LEAs with records of results from previous field trials, and because comparisons across LEAs using the same projects during a given year could readily be generated.

Because the non-incremental alternatives lie farther from SED's present programs, greater uncertainty is present in estimates of their consequences. The description of LEA processes, developed in Task 2, however, will provide a preliminary basis for making the predictions needed for the trade-off analysis.

The results of the above three tasks of Phase I will be reviewed and consolidated in Task 4 into formal recommendations concerning the program SED should adopt as potentially the most effective in raising the quality of project planning, management, and evaluation of projects funded under Urban Education and Title I, subject to constraints of costs and feasibility.

3. Phase II--Program Design

Determining what should be done is only the first step. The next is to determine how to do it. Phase II will address the second question. A formal review by SED will be conducted between the two phases to insure the relevance of the design work in Phase II. The description of Phase II must be more tentative than the description of the first phase because Phase II depends on the results of the first phase for its substance.

Phase II will consist of five tasks. Task 1 will begin with the reduction of the most promising SED program alternative(s), as determined by the review of Phase I, to its (their) component activities and will conclude with the design of a system for executing and managing these activities.

Task 2 will consist of the development of a broad strategy for implementing the system designed in Task 1. It will focus on staffing, costs, time-phasing, likely contingencies, and how these contingencies may be dealt with.

Task 3 will consist of a sensitivity analysis, a challenging of the assumptions, conclusions, and concepts incorporated into the system design and the implementation strategy. This analysis will serve to identify major uncertainties and risks so they may be dealt with appropriately. These uncertainties will be divided into those which prohibit

implementation and those which can be coped with through modifications in the program design, the implementation strategy, or the iterative operation of the program after implementation.

Task 4 will be the generation of a program development plan which is more specific than the implementation strategy drafted in Task 2. It will contain detailed plans (tactics) for implementing those parts of the program about which there is low uncertainty and plans for reducing any uncertainties which might inhibit current implementation. Plans for reducing uncertainty might include field tests, further analysis, etc.

Task 5 will consist of the drafting of any policy position papers deemed necessary to support the Commissioner and the Regents in taking action based on the results of the study. Examples of cases where such papers might be useful include the relaxation of constraints (e.g., laws) which inhibit SED's effectiveness, and the adaptation of the results of Phase I to the more general fields of categorical funding programs and compensatory education.

4. Summary

The approach described above provides:

(a) the analysis needed to determine what program SED should conduct to meet its immediate purpose of improving

the quality of planning, management, and evaluation of categorically funded projects to aid the disadvantaged;

(b) the program design necessary to manage and execute this program; and

(c) the plans and position papers needed to reduce any uncertainty regarding the program to tolerable levels and to implement the program design in an orderly and effective fashion.

II. Program Design Research: Effectiveness Measures in Reading and for School Operations

A. Background

OLRP, in the course of its work on the development of an accountability model for education in the State of New York, has analyzed norm-referenced, standardized test methodology to determine whether measures based on this approach are useful in constructing an accountability model. The results of OLRP's analysis are as follows.*

> The demand for accountability has focused an increasing amount of attention on ways of measuring the outcomes of education. This concern can be reflected in the considerable amount of debate in the past several years over the types of instruments that should be used to measure educational outcomes. Much of the debate has centered on the efficacy of norm-referenced testing with respect to its use as a tool for educational programming and evaluation.
>
> The primary conclusion that has been reached as a result of the analysis is that while these tests may serve useful purposes (if we make several stringent assumptions with such use), they cannot effectively contribute to any accountability system where the emphasis is placed on measuring student progress relative to educational treatment.

---

* Freeborne, G. Staff Paper on "Refining the Department's Policy on Racial Balance." Attachment B: Norm-referenced Testing. New York State Education Department, September 5, 1972.

This conclusion is based almost exclusively upon methodological considerations. Thus, what is being said is that even quality norm-referenced tests, because of the way they are constructed, cannot serve as the foundation of an accountability system nor as useful tools for educational evaluation. If two other considerations in addition to methodological considerations are included, namely the operational problems involved in constructing a quality norm-referenced test and the problems of interpretation of norm-referenced test results (particularly in the hands of persons not skilled in norm-referenced testing methodology), the evidence tends to indicate that these tests do much more harm than good.

To answer why these tests are not useful in any accountability framework, it is necessary to examine the major purposes of the tests, and given the purposes, examine how the tests are developed to achieve those purposes. The expressed intent of norm-referenced tests is to differentiate among students. The entire methodological framework is based upon developing an instrument which provides for a maximum amount of differentiation between individual scores. Thus, the guiding principle is one of relative comparison on individual to individual. Any individual scores from a norm-referenced test make sense only in relation to some other individual scores.

A second important principle is that you not only have differences in individual scores, but that you have wide variation in differences. Such wide variation is necessary if individual differences are to be statistically defensible, i.e., that the differences are not a result of luck, chance, etc. and that the differences are wide enough so that there is some level of confidence in making relative comparisons.

Given the concern for identifying individual differences and the need for wide variations in differences, a methodology has evolved over the past 50 years to achieve these ends. The methodology has established systematic and empirically validated ways of devising instruments to achieve differentiation and wide variability of individual scores. The emphasis has focused on constructing tests composed of a number of items of known characteristics. The most important characteristic of

an item, if differentiation is to be achieved, is that some percentage of individuals will respond to the item incorrectly. If all (or mostly all) individuals answer an item correctly (or incorrectly), that item becomes virtually useless in differentiating among individuals. Thus the driving force is toward generation of items that are neither "too easy" nor "too hard" in order to achieve a reliable degree of differentiation. (It has been proven statistically that the greatest degree of differentiation with wide variation is achieved if all or mostly all items on a test are of average difficulty, i.e., roughly half of the responses on each item are correct.)

The way a typical norm-referenced test is developed is that a number of items are generated from a specific content area, the items are pretested on a sample of students, and information on their responses is collected. Those items which differentiate best are most useful for achieving the purposes of the test. Thus, before a test item is included on the final version of the test, it is not necessarily known how the item differentiates, or why it differentiates, but what is known is that it does differentiate. The high correlation between achievement test results with intelligence or aptitude test results have led an increasing number of analysts to conclude that standardized achievement tests are not as adequate a measure of educational outcomes as they are of general mental abilities.

Norm-referenced testing instruments are poor choices for individual or program evaluation precisely because of the emphasis on differentiation and relative comparison of students. With an emphasis on differentiation, questions which all or most children can successfully answer are eliminated, or reduced to an absolute minimum. Thus, if knowledge bases of children expand (as could reasonably be expected over given time intervals), little or none of this would be perceived in an examination of norm-referenced testing results because the tendency is to eliminate common base information known by all children. The second, and even more critical aspect, is the masking of group growth because of the emphasis on relative comparison. A child is progressing "normally" if he

maintains the same relative rank on test results. Group progress (meaning improvement in scores) on the lower end of the test results only tends to raise average performance for all test results, and over time, moves the average higher. If the effects of this phenomena are combined with the effects of eliminating questions which are "too easy," it becomes fairly obvious that measurement of individual or group progress, particularly for evaluation purposes, is virtually meaningless.

A growth score is the difference between two individual scores either on the same test, equivalent forms of the same test, or comparable tests. Serious methodological problems exist in using any norm-referenced testing instrument for measuring growth or "progress," either on an individual or group basis. Testing experts have recognized the limitations in measuring growth with norm-referenced test methodology, and generally caution against the use of the tests for such purposes. The primary problem relates to an individual score and the error of measurement associated with that score. In computing individual growth scores, the error of measurement of the growth score is the sum of the errors of measurement of the two individual scores. Thus if growth scores are to be obtained which are statistically meaningful and not subject to a large error of measurement, it is necessary that such difference scores be quite large. However, when it is considered how the test is developed, it would be highly unusual to expect such large score differentials. What has been found in practice is that growth scores derived from norm-referenced tests are characterized by their extremely low reliabilities.

Significant methodological problems are also encountered in the development of norm-referenced tests with respect to introduction of cultural biases into the tests, and with respect to the adequacy of sampling the domain of knowledge to be tested. Cultural differences are normally associated with different language patterns and thought processes. To develop a standardized test which adequately takes into account such differences is impossible. What occurs is that the test is developed and normed primarily on the characteristics of the predominant culture. Any interpretations of test results of individuals who are not part of the predominant culture are suspect.

> Adequacy of sampling the domain of knowledge to be tested involves content and behavioral considerations. It is not only important to test what a student knows (content) but also what he is able to do with what he knows (behavior). The adequacy of sampling from a content domain is not as much a methodological problem as it is an operational problem. Sampling of the behavioral domain, however, is a methodological problem for norm-referenced tests because many of the important educational outcomes desired are either not clearly specified or are not susceptible to measurement at this time. Thus, norm-referenced tests have no utility in assessing or evaluating the degree of creativity in individuals, interests, positive (or negative) attitudes, values, and appreciation of art, music, culture, etc., all of which are important educational goals. Lacking the capability of measuring these variables, norm-referenced testing instruments are forced to rely heavily on content considerations. This becomes problematic if emphasis on improvement of achievement scores comes at the expense of other important, non-measurable (at this time) goals.

It is apparent from the OLRP review that norm-referenced test methodology will not yield a good measure of what students have learned from a sequence of educational programs. Furthermore, the methodology, as indicated, has limited utility in measuring attitudes and other noncognitive outcomes.

The conclusion that norm-referenced tests are not useful in any accountability framework does not mean that an appropriate set of accountability measures, or even an appropriate set of objectives for such measures, can be specified at present. Objectives for accountability have not been

defined for two reasons. First, no view or model of accountability has ever been generally adopted. Second, the state-of-the art in education does not permit a demonstration that meeting one sequence of instrumental objectives in grades K through 12 will enable students to meet desirable terminal objectives better than another sequence of instrumental objectives. In fact, taking reading as an example, the present state-of-the art in education does not enable us to know what terminal objectives (e.g., the ability to read at a certain speed and with a specified level of comprehension) to set at all.

As a consequence of such gaps in the state-of-the art, OLRP has decided to undertake R&D efforts which will yield a preliminary set of alternative accountability measures. Although present cost constraints limit the number and extent of development of these preliminary measures, even a limited set will be useful to help to define precisely the standards which are to be met by educators.

In the following sections, three different accountability measures are proposed. In accord with OLRP aims, these measures are designed to address different aspects of accountability. The first, a reading effectiveness measure, is designed to focus on the outcome aspect of accountability, i.e., on measuring how well students can read in terms that

will have meaning to reasonably educated laymen. The second measure is designed to focus on an operational aspect of accountability by determining how well school districts carry out the necessary steps to create situations that enhance learning by students. The third measure focuses on another operational aspect: the extent to which school districts have established an environment which students positively value. The latter two measures, in effect, pose questions: should schools be accountable for communicating to students what is expected of them in order that they can profit from daily learning activities; and should schools be accountable for convincing students that schools provide a positively valued service to students?

The three proposed measures were selected after extensive discussions with OLRP staff. These discussions led not only to the selection of the three measures from a set which included other candidates, but also to a mix of research tasks making up an efficient R&D program for RRI. For example, as will become clear in a subsequent discussion, the work on the reading effectiveness measure required inventiveness. Tasks requiring inventiveness are more difficult to program than straightforward developmental efforts, and progress is rarely proportional to the number of people assigned to such tasks. Therefore, rather than massing

manpower on tasks requiring inventiveness in order to satisfy SED's wish to obtain one measure (e.g., reading effectiveness) before another, it was OLRP's judgment to select a set of measures and tasks to optimally harness RRI's R&D capability. Furthermore, OLRP decided that it was preferable to take three measures through to the design concept stage* during the proposed R&D program in order to provide OLRP with a basis for structuring SED decision-making processes, rather than to attempt to take one measure completely through all developmental stages.

The proposed measures also have secondary requirements which do not directly relate to educational accountability. These requirements will be discussed as the approach to the development of each proposed measure is reviewed. They have been included to improve the utility of the measures in meeting the immediate needs of managers in education.

B. Proposed Effectiveness Measures

1. Reading Effectiveness Measure(s)

(a) Background and functional requirements. Beginning with the Coleman report, a number of recent works have concluded that education (as characterized by input variables)

---

* The meaning of "design concept stage" for each measure will be made clear subsequently in this proposal.

is not strongly related to outcomes (e.g., reading achievement) if the effects of such variables as socioeconomic class and IQ are taken into account or controlled. Yet the authors of these works acknowledge that learning how to read and how to perform basic arithmetic operations are among the capabilities largely acquired from educational programs. Furthermore, they also acknowledge that mastery of some subjects such as arithmetic must take place predominately in the schools. Therefore it is important to understand why input-output studies have not shown that educational processes result in cognitive and noncognitive development.

Part of the problem may lie in the fact that input and output measures are frequently correlated with one another because they are constructed from a common, norm-referenced methodology. For example, both IQ and standard reading tests share a common test-construction methodology. Therefore, the scores of these tests may be highly correlated not because the phenomena these tests seek to measure are empirically related, but because both kinds of tests are actually (although unintentionally) measuring the same or nearly the same capabilities or capacities.

If the method correlation between measures of input and output variables is high, it follows that a limited amount of variance remains to be accounted for by other

variables, e.g., educational programs. For example, after IQ and reading achievement scores are correlated, there is a limited amount of variance left. If this limited amount of variance is too small to attribute to variables such as school programs, then it is no surprise that many studies have concluded that school programs do not strongly affect educational outcomes.

There may be other reasons why norm-referenced tests are inadequate measures of educational accountability. Let us assume that ten years of formal education raises the reading abilities of the members of a group by some measurable amount. Norm-referenced tests probably would never detect this change in ability. Because the tests are periodically renormed, the worst-performing student always defines the worst score. Therefore, the distribution of scores always looks the same, even though the worst student might be able to read a tax form after ten years of education while the worst student before ten years of education might barely be able to read traffic signs. Thus norm-referenced tests cannot reveal how much better or worse off people are today by comparison with any earlier time. And yet, it is well known that in the United States today, the level of literacy is substantially higher than it was in the 1920s.

The above discussion strongly suggests that the State needs new and different measures of effectiveness if it is going to account to the public for its activities. In the area of reading effectiveness, the new measure must satisfy the following functional requirements:

(a.1) It must determine how well individuals can read relative to the standards that define adult reading competence;

(a.2) It must be able to measure sufficiently fine distinctions in reading achievement to allow for the determination of student progress from grade 2 to grade 12. (This contrasts with available literacy measures which are too crude for monitoring the performance of students in the primary grades;)

(a.3) If possible, the measure should enable the SED to set minimum quantitative and qualitative objectives (standards) so that school systems can be held accountable;

(a.4) If possible, the measure should increase the validity of decisions affecting manpower resource allocation and the value of cost-benefit and other input-output studies, whether they are executed for po icy-related or operational reasons.

(b) Alternative approaches for developing standards of reading competence.

(b.1) Expectations defining adult reading competence. In order to develop a reading effectiveness measure, the terminal reading capabilities expected must be defined. There are several approaches which could be used to formulate such standards:

- Specify a population of adult written materials and sample from this population a set which is then subjected to analysis and, ultimately, submitted to a panel of judges to define the required standards. The standards could be stated in the following form: "if certain specified conditions are met, an adult should be able to exhibit certain specified behaviors";

- Use any of a variety of methods (e.g., the Delphi Techniques*) to produce a consensus on the standards;

- Investigate theoretical models of language to see if they can be successfully used for the purposes of defining the standards.

---

* Weaver, T.W. Developing an accountable consensus of goals: The Delphi Technique. In Browder, L.H., Jr. (Ed.) Emerging problems of administrative accountability. Berkeley: McCutchen Corp., 1971.

The National Assessment Program appears to have elected the first alternative. The approach is expensive, somewhat arbitrary (experts with established points of view define the standards), the open-ended (language changes as society changes, and therefore it is necessary to re-examine standards periodically).

Problems are also associated with the second approach. Attempts to reach a concensus on any go in education have produced useful results only when methods for lowering the level of abstraction in discourse have been introduced. Thus, for example, if the work required under the first alternative were accomplished, and if the people of New York State were given an opportunity to study the results, then something might be expected from using time-consuming and expensive consensus-producing processes.*

The last alternative has never (to RRI's knowledge) been tried. Sufficient information exists concerning at least some aspects of language to recommend that formal models be examined to determine if they provide

---

* Although RRI does not recommend holding consensus-validating sessions throughout the State of New York, the people in the end have the prerogative to set and validate standards. It would be desirable to use methods other than consensus-contriving techniques for this purpose.

enough information to enable the formulation of the required standards of measurement. This approach would anchor performance standards to hard data. In addition, the third alternative has other advantages. First, it is probably the least costly approach. Second, the approach may yield estimation methods to enable linguistic changes to be forecasted.* Third, and most important, the approach holds the promise of empirically testing assumptions concerning the importance of different sequences of expectations set for grades 2 through 12 to the achievement of adult competence

---

* All language has a history. The literature shows that some aspects of language change more rapidly than others. Some of these changes are clearly of limited interest in setting expectations. For example, words come and go in English (as in every other language) as a function of our major preoccupations. The word "flack" for example was introduced during World War II as an abbreviation of a German expression for an antiaircraft gun firing fragmentation shells. Many other archaic terms are with us still, and other words have long since passed out of common English use.

Some aspects of language change more slowly. It is precisely these aspects, however, which should be the subject of projections. More specifically, we need to determine whether materials written for the ordinary citizen have increased or decreased in "readability" over the last thirty years, and to measure the rate of increase or decrease in "readability." This work may lead to assessments of what the adult world will expect some fifteen years from now of a five-year old currently entering kindergarten. On the hypothesis that we work out methods for measuring the "readability" of written materials, then it should be possible to make such projections.

in reading.* This last application of measures of reading effectiveness is discussed further in the next section.

(b.2) Setting instrumental objectives. A measure of reading effectiveness must determine not only the extent to which an individual has achieved adult competence, but also the extent to which a student, at a given developmental stage, is progressing toward the expectations set by adults in the socialization process. Any of the alternative approaches outlined in (b.1) could be used to set the standards for each grade. However, the state-of-the art is such that no expert can be sure that, by acquiring capabilities X_1, X_2, . . ., X_n in Grade 2, capabilities Y_1, Y_2, . . ., Y_n in Grade 3, and so on, a student will eventually acquire adult competence in reading. The state-of-the art does not even permit the estimation of the probability of arriving at any level of reading competence if specified steps are followed.

Any consensus reached by adults using the Delphi or any other technique (even if it were efficiently arrived at for each grade) would simply be a

---

* This work is beyond the scope of the proposed R&D research but it is an important component of a larger effort leading to reading curriculum validation.

consensus which could be worthwhile or worthless depending upon how well adults from all walks of life intuitively understand human development. The third alternative, while it does not yield direct knowledge of the standards which should be set in grades 2 through 12, has the prospect of at least describing present expectations, and, in addition, of setting the stage for future research to show how the acquisition of reading skills can be simplified.

The discussion of alternative approaches, both with respect to setting expectations (standards) of adult competence as discussed in subsection (b.1), and with respect to setting instrumental objectives for each of the primary and secondary grades as discussed in subsection (b.2), suggests that the third approach--the use of formal models of language to develop the standards of reading competence at all levels--should be pursued.

c. Method of procedure. Discussions were held with OLRP concerning the trade-offs among alternative approaches to developing standards of adult competence in reading, and the problems associated with setting standards for each of the grades 2 through 12. OLRP accepted RRI's recommendation that an R&D instrumentation effort be undertaken to determine if formal models of language could provide a basis for developing both standards of reading

competence and new reading effectiveness measures.

On the basis of an incomplete literature search, it appears that an experimentally confirmed mathematical (lognormal) model of words (tokens) and word types has been formulated for the written, adult language. Furthermore, Carroll et al.* have shown that the tokens and types contained in written materials used in grades 3 through 9 also can be described by lognormal distributions.

If the lognormal model adequately describes the distribution of tokens and types in written English, then a quantitative description of written materials in terms of their level (i.e., grades 3 through 12 and adult) is at hand. Furthermore, it should be possible to quantify the relationships among the written materials used in different grades and adult written materials. Finally, provided that samples of words are properly drawn, and provided that students know the meanings of these words, it should be possible to assess quantitatively the extent to which a student has progressed toward mastering adult vocabulary in each of the grades in which reading is taught.

The existence of such a model raises many possibilities for the identification of instrumental

---

* Carroll, J.B., Davies, P. & Richman, B. Word frequency book. New York: Houghton Mifflin, 1971.

objectives for reading programs. With such a model, samples of written text used in various grades could be analyzed and significant differences identified. From these differences, the instrumental objectives which are currently governing instruction in each grade could be inferred. Suppose, for example, that analysis showed that, relative to adult English text, the most common words appeared more frequently in the lower than higher grades, that moderately common words appeared less often in the lower grades than in the middle or upper grades, and that the rarest words appeared only in the upper grades. From this information, a variety of instrumental objectives could be inferred: concentrate almost exclusively on teaching common words in the earliest grades; teach moderately common words in the middle grades and reduce the emphasis on reading the most common words; and introduce rare English words in the upper grades while sustaining, but at a reduced priority level, an interest in teaching the common and moderately common words.

Once the instrumental objectives have been identified, it should be possible to measure the extent to which students have met the expectations embodied in the materials used in the various grades. Furthermore, if samples of adult written materials (newspapers, tax forms, etc.) were related to the model describing the full adult

written language, then a rational basis for defining the capabilities for competent adult performance in reading should be feasible.

On the hypothesis that the R&D program is successful and RRI is able to properly sample word types for different purposes (to measure whether the expectations on students in each grade are being fulfilled and to measure student progress towards becoming a reader of adult-level text), it should be possible to devise test score reports that are understandable to parents as well as to other adults. For example, we might be able to report that student X has learned the meaning of nearly all of the most common English words, most of the next most common words, etc. and, in the aggregate, has acquired Z% of adult competence. Since Carroll* has shown that adults can accurately estimate the relative frequency of English words, such scores should be intuitively meaningful to parents and others. This procedure would represent an advance over the present practice of reporting test results in technical jargon, e.g., national or regional grade norms, which many educators, no less laymen, don't understand.

---

* Carroll, J.B. Measurement properties of subjective magnitude estimates of word frequency. Princeton, N.J.: Educational Testing Service, 1971.

Formally describing the distribution of types and tokens in English may not describe English sufficiently well for our purposes. Undoubtedly educators would want to test to determine whether students comprehended passages of English text. Adults expect that children will improve their grasp of the meaning of written materials as they grow. More particularly, adults expect that learners will understand more complex explanations as the learners develop. Furthermore, as learners grow, they are expected to acquire the ability to infer the intent of a writer from his selective use of linguistic terms and forms, to be sensitive to alternative inferences which may be drawn, and to be aware of the variety of emotional effects which the writing might have on other readers.

There is no known way to completely describe all the meanings of written materials. However, all that is needed at present is the capability to scale the difficulty, or so-called "readability," of passages of text. Such scaling is feasible.

It should be feasible to characterize the "readability" of materials used at different grade levels as well as samples of adult materials. If a passage of English text were given a score on a "readability" scale, then it should be possible to draw samples of English text

of known "readability" for the purposes of testing comprehension. For example, if a student scores well on comprehension for the most "readable" text, less well for the next most "readable" text, and so on, then it should be possible to give the student an interpretable score. Furthermore, since one of the criteria used in the literature for scaling the "readability" of English text is type and token content, it may be feasible to translate scores on reading comprehension into mastery scores based upon the significance of knowing the meanings of different words with different probabilities of occurrence.

Syntax is another aspect of language which might have to be taken into account in formulating a reading effectiveness measure. It is well known that, as children grow, they are able to understand increasingly complex syntactical forms. At this time, however, there are many issues to be resolved concerning the best method for precisely describing the syntactical complexity of English text. Certainly a variety of linguistic characteristics (e.g., sentence length) could be measured. Whether any such indicators, or some combination of them, can adequately describe the increasingly complex set of expectations imposed on learners progressing from grades k through 12, is an open question. In addition, at this stage of

the research, much uncertainty surrounds the need to develop a separate scale of syntactical complexity or a separate test for mastery of syntax. A review of the literature might show that syntax is so hopelessly confounded with vocabulary and with "readability" that a separate test for mastery of syntax would be unnecessary to obtain valid reading effectiveness scores. Indeed, since the various aspects of language seem to be interrelated, it remains to be seen whether more than one aspect needs to be sampled for mastery in order to formulate an adequate effectiveness measure. On the other hand, people might be more inclined to accept test results if they felt that all aspects of language mastery were tested. Furthermore, having several independent ways of estimating progress towards becoming a competent adult has its advantages from a construct validity point of view.

2. Orientation Towards Learning: A Measure of the Effectiveness of School Operations

(a) Background. In colleges, formal courses and informal settings bringing together students and faculty are extensively used in an attempt to orient students to life in the institution and to the content and boundaries of various academic disciplines. In any educational environment, attempts are made to structure all learning situations by using some means of communication (usually speech and

written text) to explain what is expected of students, both with respect to their learning activities (calculating, reading, role-playing, etc.) and with respect to the aims, or expected outcomes of these activities. The capability for setting up successful learning situations depends on the ability of teachers to communicate instructions to students. Since this ability is an invariant expectation on teachers over the entire student age range and over all modes of instruction, OLRP and RRI have considered the development of a measure to determine the extent to which this expectation is being fulfilled, especially in the primary grades. OLRP has decided, as indicated previously, that it would be desirable to develop a prototype measure in order to be able to put the question in less abstract form to SED decision-makers. Therefore, RRI has agreed to propose the development of such a prototype measure.

(b) Method of procedure. One way of determining whether teachers are successful in setting up learning activities is to test whether students understand key instructions, for example, the concept "given." An understanding of such key concepts by students is a critical pre-condition for successfully setting up learning situations. Differentiating between one instruction, "read," and another instruction, "write," is also important to the success of

learning activities. Finally, insuring that all students share the meaning of key concepts and instructions is important for two reasons. First, instruction in school situations are most often given to a group of students, and second, there is a heavy reliance in educational institutions on students orienting each other so that group learning is facilitated.

A prototype measure may be formulated by sampling a set of key concepts related to the transmission of instructions to students in certain grades (say 2 through 12). If a synonym for each key concept in the sample were included in a set, then students could be asked to judge the degree of similarity between all pairs of concepts in the set. If a particular pair of words meant the same thing, the similarity judgment should be at its highest value. If the pair of words had little similarity, the similarity judgments should be minimal.*

Analyses of the judgments should permit a determination of the degree to which students comprehend important concepts (i.e., can identify the synonyms) and understand significant differences among concepts. In

---

* It is likely that the form of the test and the instructions will be based on the People Test (see "Guidelines for the evaluation of desegregation programs in school districts," RRI, September, 1972.

addition, it should be possible to derive a measure of the extent to which each student shares his similarity judgments with every other student. RRI believes that recent developments in multi-dimensional scaling techniques may be useful in developing the required measure.

(c) Other applications. The proposed measure has several other possible applications. First, it could be employed as a moderator variable in policy-related research. For example, if class size is reduced, some teachers may be better able to get students to understand and share instruction-related concepts than other teachers. This phenomenon might explain the weak relationships reported in the literature between class size and the cognitive outcomes of educational programs. Thus, the measure could be used (along with others) in investigations designed to determine optimal class size under different circumstances, e.g., different classroom mixes of racial and socioeconomic heterogeneity. Second, the measure might prove useful in research directed toward understanding why some teachers are more effective than others. (For example, good teachers may spend a good deal more time early in the semester getting students oriented towards what is expected of them.) Third, the measure could be useful in accounting for observed differential effectiveness among educational programs designed to meet the same objectives.

3. Valuing Learning and School: A Measure of the Effectiveness of School Operations

(a) Background. An invariant expectation on all educators is that they should structure an environment which is positively valued by students. The rationale for this expectation is based both on research findings and on observations made over the sweep of Western Civilization. Among the research findings are those which demonstrate a causal link between achievement and liking a teacher, or a field of study, or school in general. Among the axioms of educators are such statements as: "If rapport is not established between learners and teachers, little will happen of any value"; or "Learning induces anxiety since we have to venture and make mistakes. If a student believes that he will be negatively valued for making mistakes, he is apt not to volunteer to learn. Hence he will, to maintain self esteem, reject or ignore education." Therefore OLRP, as indicated previously, has requested that RRI develop a prototypical measure of student attitudes towards educational environments. Again, the rationale is to develop a measure which could be used by decision-makers as another measure of accountability.

(b) Method of procedure. RRI proposes to further develop its Test of the Special Meanings of Words,

which was designed and field-tested in previous work performed for SED. A description of this test follows.[†]

> Test format and administration. Students (grades 5-12) are asked to rate each concept against a set of bipolar adjectives. One set of eleven adjective scales is used for rating human/animate concepts and another set of eleven for abstract/nonanimate concepts. Adjectives of obvious social desirability are excluded from the animate adjective group and descriptive or clearly denotative adjectives are excluded from the abstract/nonanimate group. The difficulty of the vocabulary is controlled so that all words in the test should be familiar to most upper third graders. Concepts and adjectives are either A or AA words in Thorndike-Lorge word book,* or appear on the Dale-Chall** or Stone *** easy word lists.
>
> Since grade school children appear to work better and with greater sensitivity on a five-point scale,**** such a scale has been used in the Test of the Special Meanings of Words rather than the more frequently used seven-point scale. The center of the scale constitutes a neutral

---

[†] "Effectiveness of racial balance programs: Final report for Phase IIA, volume I," RRI, September, 1971.

* Thorndike, E. & Lorge, I. The teacher's word book of 30,000 words. New York: Teachers College Press, 1944.

** Dale, E. & Chall, J. A formula for predicting readability: Instructions. Educational Research Bulletin, 1948, 27(2), 11-28.

*** Stone, C. Measuring difficulty of primary reading material: A constructive criticism of Spache's measure. Elementary School Journal, 1956, 57, 36-41.

**** Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. The measurement of meaning. Urbana: University of Illinois Press, 1957.

position, and the ends constitute the polar extremes of the adjective pair. All scale points are clearly labelled (anchored) as to meaning, following the suggestion of McNeil and Phillips.* Students are asked to decide which of the five scale points of the adjective pair comes closest to describing how the concept being rated "feels" to them, and to place a checkmark at that point along the adjective scale.

Factor analysis. Factor analytic studies of the Test of the Special Meanings of Words were carried out to determine whether factor structures were sufficiently comparable for black and white students at different levels of racial heterogeneity to permit valid comparisons between groups.

Scales by students matrices were obtained by summing over concepts. A separate matrix was obtained for black and for white students at each of four levels of racial heterogeneity. Each matrix was then intercorrelated to produce eight symmetric intercorrelation matrices. A principal components factor analysis was performed on each correlation matrix. The resulting factors were rotated to simple structure using a varimax rotation scheme. The factor structures were then compared. Since the first three factors accounted for over 80% of the common variance, the first three factors were used for comparisons between groups. Coefficients of congruence** were calculated to compare the factor structures of the eight groups. The coefficients were all exceedingly high, indicating that the groups used sufficiently similar frames of reference in rating the concepts to allow comparisons between blacks and whites and between students in classes of different racial compositions.

Validity studies. There are many studies in the research literature which document the

---

* McNeil, K.A. & Phillips, B.N. Scholastic nature of responses to the environment in selected subcultures. Journal of Educational Psychology, 1969, 60(2), 79-85.

** Wrigley, C. & Neuhaus, J.O. The matching of two sets of factors. American Psychologist, 1955, 10, 418-419.

predictive and construct validity of the semantic differential (SD) technique (e.g., Osgood et al.* and Reeves**). Moreover, it has been demonstrated that the evaluative dimension of the SD is an attitude measure, e.g., Likert, Thurstone, and Guttman attitude scales are highly correlated with the evaluative factor of the SD. Thus semantic differential data include attitude as a component of the "meaning" they index.

Since the validity of semantic differential techniques as attitude measures has been well documented, construct validation analyses for the Test of the Special Meanings of Words have been limited to demonstrating that the four subtests of the instrument measure the psychological constructs they were designed to assess. Thus validation analyses have been designed to show that indices derived from selected concepts differentiate between groups believed to differ with respect to sense of fate control, social class, self-concept and attitudes towards school. Preliminary studies comparing white and black students in the test-bed sample are reported in this section. These black-white comparisons make use of the statistic D, a generalized distance function which is an index of differences in connotative meaning between concepts.*** For each student who took the Test of the Special Meanings of Words, a D score was calculated for each of several key concept pairs. Analyses were then carried out to compare the D scores of white students with the D scores of black students.

---

* Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. The measurement of meaning. Urbana: University of Illinois Press, 1957.

** Reeves, M.P. An application of the semantic differential to the Thematic Aperception Test material. Unpublished doctoral dissertation, University of Illinois, 1954.

*** Cronbach, L.J. & Gleser, G.C. Assessing similarity between profiles. Psychological Bulletin, 1953, 50, 456-473.

Osgood, C.E. & Suci, G.J. A measure of relation determined by both mean difference and profile information. Psychological Bulletin, 1952, 49, 251-262.

Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. The measurement of meaning. Urbana: University of Illinois Press, 1957.

The Test of the Special Meanings of Words was originally developed to test hypotheses concerning some of the effects of school desegregation. (Other purposes are indicated in RRI's report. The instrument proved to be extremely sensitive, see pp. 50-51 of the report.)

Analyses of the Test of the Special Meanings of Words provides a basis for developing an efficient and more polished measure of attitude, in which the semantic differential technique will be used to measure the extent to which students share and positively value school-related concepts. Thus, it should fulfill the requirements of an accountability measure.

(c) Other applications. In RRI's report, "Effectiveness of racial balance programs: Final report for Phase IIA, volume I," several other applications of the instruments are described. In the work proposed here, one additional application will be attempted simply because the proposed accountability instrument can be constructed efficiently to meet an additional objective.

RRI proposes to develop a score (perhaps a score based upon the D statistic) to identify students who do not share their evaluations of components of school environments with others. RRI will develop the theory of such a measure for potential use in calling to the attention

of school administrators those students who negatively value school concepts. Such students might deviate sufficiently from others to be candidates for special courses of action. For example, the measure might identify candidates for a drop-out prevention program. Considering the number of drop-outs in urban areas (in New York City, 60% of the Spanish surnamed students eventually drop-out, and some 50% of the black students drop-out), it might be very worthwhile to detect potential drop-outs so that something could be done for them before rather than after they actually leave school.

III. Work Plan

This work plan consists of two parts. Part A contains tasks which follow from Section I of this proposal. Part B contains tasks which follow from Section II of this proposal.

A. A Program Analysis of Categorical Funding

1. The goal of this effort is to determine how SED can improve planning, management, and evaluation of projects funded under Title I and Urban Education. The objectives of the two phases which will fulfill this goal are as follows:

(a) Phase I: To conduct the program analyses necessary to recommend to SED the program which holds the greatest promise for improving planning, management, and evaluation of categorically funded projects to aid the disadvantaged.

(b) Phase II: To prepare the designs, plans, and position papers that will be necessary for the orderly and effective implementation by SED of the recommended programs.

2. Work Tasks

Riverside Research Institute (RRI) will apply its best efforts to carry out the Phase I and Phase II tasks enumerated below in conjunction with staff from the Office of Long Range Planning.

(a) Phase I: Program Analysis.

Task (A.1): A description and evaluation of SED's present operations under ESEA Title I and the New York State Urban Education Act.

Task (A.2): A trade-off analysis among incrementally different alternatives to SED's present Title I and Urban Education programs.

Task (A.3): A trade-off analysis among non-incrementally different alternatives to SED's present Title I and Urban Education programs.

Task (A.4): The consolidation and review of the findings of Tasks 1, 2 and 3, and the drafting of a recommendation to SED identifying the most promising program(s) for implementation.

(b) Phase II: Program Design.

Task (A.5): The determination of the component activities of the most promising program, and the design of a system for the execution and management of these activities.

Task (A.6): The development of a general strategy for implementing the program.

Task (A.7): A sensitivity analysis to identify critical uncertainties in the program design and implementation strategy developed above.

Task (A.8): The generation of a program development plan for reducing critical uncertainties and implementing the program.

Task (A.9): The drafting of any policy position papers deemed necessary to support the implementation of the recommended program. The drafting of position papers will be undertaken by mutual agreement between RRI and SED.

3. Management Plan

The work described in the proposed research will be monitored for SED by Mr. Gerald Freeborne, Assistant Commissioner for Long Range Planning. It will be directed for RRI by Dr. Ben Josephson, Jr., and will be managed by Mr. J. P. Brashear.

(a) Reports and schedule.

The proposed project will run from July 1, 1972 to June 30, 1973.

Bi-monthly progress letters will be transmitted to Mr. Freeborne. These will be supplemented by informal working papers from time to time, as mutually agreed upon by SED and RRI.

Formal written reports, supplemented by oral presentations if desired by SED, will be submitted at the completion of each of the two phases described in Section A.1. of this Work Plan. These reports will contain the results and recommendations of the work in each phase.

A review will be conducted under Mr. Freeborne's direction at the completion of Phase I in order to provide the guidance necessary to ensure the relevance of the design work in Phase II.

(b) Adjustments to the work plan.

Any modification in this Work Plan will be by mutual agreement of Mr. Freeborne for SED and Dr. Josephson for RRI.

B. Program Design Research: Effectiveness Measures in Reading and for School Operations

1. Project Goal and Objectives

The primary goal of this research is to develop the design concepts for one or more measures of reading effectiveness for the primary and secondary grades. The secondary goal of this research is to develop design concepts for two prototype measures of accountability that will establish how well school districts (and subordinate units) fulfill socially shared expectations.

2. Work Tasks

RRI will apply its best efforts to execute the following tasks.

(a) The development of design concepts for a reading effectiveness measure (or measures).

Task (B.1): Fundamental Research

[i] Investigate and, if necessary, extend existing statistical models of the corpus of written English words and word types in order to permit an evaluation of the extent to which appropriate standards of reading capability (for the adult level and for different grade levels) can be established from the models, and a determination of how reading effectiveness measures based on these models might be constructed.

[ii] Investigate the validity of existing methods of describing the syntactical complexity of written English, evaluate the need to include syntactical elements in defining appropriate standards of reading capability (for the adult level and for different grade levels), and determine whether (and, if so, how) reading effectiveness measures based on syntactical models can be constructed.

[iii] Investigate and, if necessary, further develop methods for scaling the difficulty or "readability" of English prose in order to permit an evaluation of the utility of such methods in defining appropriate standards of reading comprehension (for the adult level and for different grade levels), and a determination of how effectiveness measures based on "readability" might be constructed to meet various objectives.

Task (B.2): Design Research

[i] Define detailed functional specifications for reading effectiveness measures. In particular, RRI will apply its best efforts:

- To determine how minimum standards of reading capability for each grade level and for the adult (terminal) level could be formulated;
- To determine how qualitative, incremental objectives for the development of progressively higher reading capability levels at different stages (grades) of the educational process could be constructed;
- To determine how progress towards adult reading capability can be defined operationally;
- To determine how the extent to which districts fulfill grade-level objectives and produce group gain in reading competence should be measured;
- To determine how the validity of decisions affecting manpower resource allocation, and the value of cost-benefit and other operational or policy related studies, can be increased through the use of reading effectiveness measures.

[ii] Define design criteria for reading effectiveness measures. In particular, RRI will consider:

- The nature of the scale required to meet the functional requirements in [i];

- The unit of measure;
- The required reliability level;
- The required face, construct and predictive validity.

[iii] Define design constraints for reading effectiveness measures. In particular, RRI will consider:

- Total time for test administration;
- Permissible unit cost;
- Required no-program bias.

Task (B.3): Test Methodology Development

Formulate design concepts for one or more tests of reading effectiveness which meet the conditions defined in Task (B.2). In particular, RRI will consider:

- How samples of words, syntactical forms, or passages of text should be drawn for purposes of test construction in order to satisfy the models investigated in Task (B.1);
- The development of appropriate test formats and prototype questions;
- The definition of scores;
- Test reliability.

Trade-off analyses will be performed to clarify problems in test construction. The analyses will trade off the functional requirements against the risks and uncertainties inherent in the models and in the methodology.

(b) The development of design concepts for two measures of the effectiveness of school operations.

Task (B.4): Develop Design Concepts for a Measure of the Extent to Which Students Have Been Properly Orientated Towards Learning

In particular, RRI will consider:

- The functional requirements for the measure;
- The design concepts for the test instrument, including the definition and analyses of scores, test formats, prototype questions, etc.

Task (B.5): Develop Design Concepts for a Measure of the Extent to Which Students Value Learning and School

In particular, RRI will consider:

- The functional requirements for the measure;
- The design concepts for the test instrument, including the definition and analyses of scores, test formats, prototype, questions, etc.

3. Management Plan

The work proposed in this Section of the Work Plan will be monitored for SED by Mr. Gerald Freebc:ne, Assistant

Commissioner for Long Range Planning. It will be directed for RRI by Dr. Ben Josephson.

(a) Reports and schedule.

The proposed project will run from July 1, 1972 to June 30, 1973. Progress letters will be transmitted on an "as needed" basis to be determined by Mr. Freeborne.

A final written report, supplemented by oral presentations (if required by SED) will be submitted following the completion of all tasks in Section IIIB.

(b) Adjustments to the work plan.

Any modifications in this Work Plan will be by mutual agreement of Mr. Freeborne for SED and Dr. Josephson for RRI.

RIVERSIDE RESEARCH INSTITUTE*

SOCIAL SYSTEMS DIVISION

KEY PERSONNEL

---

* Riverside Research Institute does not discriminate against any employee or applicant for employment because of race, creed, color or national origin and takes affirmative action to insure that its employees are afforded equal employment opportunities. Such action is taken with reference, but not limited, to: recruitment, employment, job assignment, promotion, upgrading, demotion, transfer, layoff or termination, rates of pay or other forms of compensation, and selection for training or retraining.

RRI Organization

Riverside Research Institute is a not-for-profit New York State membership corporation. RRI was formerly known as the Columbia University Electronics Research Laboratories which was formed on June 1, 1951.

The total RRI budget is currently $10.0 million per year. Of the 340 employees, 150 are professional members of the research staff consisting of engineers, applied physicists, mathematicians, programmers, research psychologists, sociologists, economists, educators, and members of other scientific and engineering disciplines.

Riverside Research Institute currently consists of four major divisions, several special facilities and appropriate technical and administrative supporting services. Current research programs at RRI include the following:

Military Systems Studies

Theoretical and evaluative studies and technological support for satellite, ballistic missile, interceptor and aircraft systems; weapons, sensors, data processors, etc., and their strategic use.

Experimental Research in Physics, Biology and Engineering

Research programs are currently being operated in optics, electro-optics, acoustics and ultrasonics, electronics, plasma engineering, biomedical engineering, mathematics, mechanical design and computer sciences.

Radar Technology

This program includes the direction of the AMRAD Measurements Program, as well as the development and implementation of advanced technology for the AMRAD radar facility (White Sands Missile Range).

Social Systems and Urban Studies Divisions

Programs pertaining to research and evaluation in urban problems are handled by groups which include Research Psychologists, Social Psychologists, Mathematicians, Statisticians, Computer Specialists, Systems Analysts, Educational Specialists and personnel in other associated scientific disciplines.

Security

The Riverside Research Institute was granted a top Secret Facility Clearance on October 21, 1968 by the Defense Contracts Administration Services (Region, New York).

Facilities

Riverside Research Institute's main facility is located on Manhattan's upper west side. Engineering offices, research laboratories, a large digital computer and various support activities occupy over 80,000 square feet of space. Riverside Research Institute also directs the operation of the AMRAD Radar Facility located at the White Sands Missile Range in New Mexico.

The computing needs of the Institute are presently met by an on-site Sigma Seven computing facility and conveniently

located IBM 360 equipment situated off-site. These facilities are utilized for a wide variety of non-routine scientific problems, standardized data reduction procedures, and assorted business applications.

The Social Systems Division of RRI maintains a fully integrated and interlinked set of computer programs for statistical data reduction and data bank maintenance. Statistical procedures currently available include analysis of variance and covariance, all correlations and linear multiple regression analysis, t-tests for independent and correlated data, factor analyses and principle components analysis, analyses of uncertainty, chi-square analyses, and multiple discriminant analysis. This interlinked set of statistical packages is fully coordinated with data display programs to obtain, for example, frequency distributions, scatter plots, and multi-dimensional contingency tables. This program package also contains a variety of data checking and manipulation features such as merging and updating matrices, reports of bad and missing data, card sequencing checks, etc.

Support facilities at the Institute include modern well-equipped machine and electronics model shops, a drafting department, reproduction and photographic facilities as well as a library of reference work, current periodicals and technical reports. In addition, the vast library facilities of Columbia University have been made available to staff members of the Institute.

Research Experience in Societal Problems

Personnel in the Social Systems and Urban Studies Divisions have had primary responsibility for the implementation of RRI programs in urban affairs, education, social systems research and projects geared to the solution of urban problems. Specific on-going programs or those recently completed are summaried below:

Feasibility Study for a Blood Management System

RRI is presently conducting this study for the Community Blood Council of Greater New York under a grant from the National Institutes of Health. The CBC is a clearing house for blood and its derivative components, serving the greater metropolitan area.

Wastage of blood due to poor management policies in various hospital blood banks can account for up to a 30 per cent loss of fresh blood due to outdating. Because information exchange between various points in the system is inadequate, one hospital blood bank may be short of a particular type or component of blood which is in oversupply at another blood bank. The relationships between donor psychology, hospital usage and other components of the system are being identified so that peaks and valleys in the supply are evened out and a better distribution system developed. Since many blood diseases do not appear for as much as 180 days following a blood transfusion, a donor-recipient link should be established to

trace sources of infection. RRI, under subcontract to CBC, has undertaken a feasibility study to determine the extent to which a computerized system can solve some of these problems.

A Study of the Emergency Reporting System in New York City

This is an effort to provide the Fire and Police Departments of New York City with information regarding the operation of a new Emergency Reporting System which will give the public a two-way voice communication link between a street location and the Fire or Police Departments. RRI will provide technological, statistical, social and psychological evaluations which will permit optimal use of the new system for both departments as well as the public. On the basis of collected information concerning telephone communications and other variables, RRI will suggest the functional specifications for the new system.

Evaluation of Science and Mathematics Curricula in Two-Year Colleges

Under a grant by the Department of Health, Education and Welfare, the present science and mathematics curricula being offered at two-year colleges in New York State were evaluated. A final report was prepared for H.E.W. which delineated the results of the evaluation study and made recommendations for changing and unifying these programs.

Social Science Research in Educational Problems

With the support of an NIH grant the researchers at RRI have developed measures of the attitudes of children toward components of their school environment including peers, teachers, and various other aspects of school life. These measures are being employed to:

1. Determine the effect of school integration on the attitudes of Negro and white children toward one another and toward adults of different races.

2. Determine the relationships between attitudes toward components of the school environment and cognitive growth.

3. Evaluate the effects of special programs designed to change attitudes towards people and towards education.

4. Determine the long-range effects of social educational policies on attitude change and cognitive growth.

Evaluation of New York State Racial Impact Programs

The New York State Education Department is currently supporting a large number of programs to create racial balance and equal educational opportunity: rental of relocatable classes, school bussing, school construction, the usage of teacher aides, school-community coordinators, remedial assistance and the like. RRI is currently engaged in assessing and evaluating the multiple effects of the SED's racial imbalance project on children, teachers, and other components of school environments. On the basis of this evaluation, RRI will make recommendations for future funding.

In addition to the funding recommendations with regard to the racial imbalance correction and equal educational opportunity project, RRI plans to provide the State Education Department with the capability to proceed independently with future program evaluation. Relevant computer software packages may be delivered along with analytic between- and within-district statistical designs that have general application to SED's program evaluation requirements.

Consulting Relationship with the New York State Education Department

RRI currently provides several types of consulting assistance to the New York State Education Department (SED). Current support provided to SED includes the following:

Consulting Efforts Designed to Coordinate Local and State Efforts to Develop a Functioning Program Planning and Evaluation (PPE) System for Occupational Education

Three PPE subsystems are under development: an evaluation-reporting subsystem, a targeting subsystem, and monitoring subsystem. The evaluation-reporting subsystem is designed to enable occupational education administrators to meet State and Federal reporting requirements and to provide the capability for assessment of program effectiveness. The monitoring subsystem provides the capability to monitor the progress of students in on-going occupational education

courses. The targeting subsystem enables the determination of the mix of occupational education programs which is "targeted" to local, regional, or statewide job market conditions.

Consulting Support to the Office of Long Range Planning

RRI provides general consulting support and technical assistance to the SED Office of Long Range Planning for the organization of State planning processes, as well as a variety of specific State planning efforts. These efforts include the coordination of State and district level program planning and budgeting, education evaluation and accountability.

Program Monitoring for School Districts

RRI has supported the development and maintenance of program monitoring systems in two school districts. Monitoring is concerned with assessing the extent to which students are meeting the curriculum objectives of programs in progress, and with the flexible allocation of educational resources to facilitate meeting the objectives.

MA-5 Training-Support Program for the Hard-Core Unemployed

The most serious problem confronting the hard-core umemployed is the absence of opportunities for gainful employment and training in tasks which require established education and skills. Experience has demonstrated that conventional training programs concerned exclusively with technical curricula

have only a very limited chance of success. Training in specific technologies poses many problems. They include overcoming the inadequate initial educational level of the trainees, and the personal difficulties that are bound to follow any intensive attempt at improvement.

Thus RRI's MA-5 program includes not only technical vocational training but also includes extensive initial orientation and counseling, job related basic education, special counseling and job coaching, supervisory and human relations training, as well as medical and dental services as required. These program elements are applied to each trainee on an individualized basis. Training is being currently offered for the following positions: computer peripheral equipment operator, electronics technician, mechanical draftsman and general machine operator. The program has a nine month duration.

Statistical Consulting and Evaluation

RRI provides data base maintenance, statistical consulting, statistical data reduction, and program evaluation services to school districts as well as to other not-for-profit organizations.

BERTRAM L. KOSLIN

Education

1950 The Johns Hopkins University
A.B. Biology
1963 University of Oklahoma
Ph.D. Psychology

Academic and Professional Experience

1970 - present
RIVERSIDE RESEARCH INSTITUTE
New York City
Vice President for Research
and Director, Social Systems Division

1968 - 1970
RIVERSIDE RESEARCH INSTITUTE
New York City
Assistant to the Vice President
for Research
and Head, Social Sciences Research
Laboratory

1964 - 1968
PRINCETON UNIVERSITY
Princeton, New Jersey
Assistant Professor

1961 - 1964
BARD COLLEGE
Annandale-on-Hudson, New York
Assistant Professor

1960 - 1961
FAIRLEIGH DICKINSON UNIVERSITY
Rutherford, New Jersey
Instructor

1958 - 1959
UNIVERSITY OF OKLAHOMA EXTENSION
Norman, Oklahoma
Instructor

1958 - 1959
UNIVERSITY OF OKLAHOMA
Norman, Oklahoma
Research Associate, Institute of
Group Relations

1956 - 1958
UNIVERSITY OF OKLAHOMA
Norman, Oklahoma
Graduate Assistant

Academic and Professional Experience (continued)

1955 - 1956	U.S. ARMY Education Advisor and Representative of the University of Maryland's extension program for the southern district of France
1953 - 1955	UNITED FUND WELFARE ORGANIZATION (UNITED SEAMEN'S SERVICE) Representative in Japan, France and Italy
1950 - 1952	THE JOHNS HOPKINS UNIVERSITY Baltimore, Maryland Research Associate

Major Consulting and Professional Positions

1968 - present	Canada Council, Social Sciences and Humanities Division
1968 - present	Riverside Research Institute, New York City
1967 - present	Woodrow Wilson Foundation, Princeton, New Jersey
1966 - 1967; 1969	Panel Member: National Science Foundation
1967 - 1968	Two Ford Foundation projects through Institute of Educational Development.
1967 - 1968	Educational Testing Service, Princeton, New Jersey
1967	The Board of Regents of the State of New York, through Institute of Educational Development. (Member of a five man study panel for the N.Y.S. Commissioner of Education on the problems of decentralizing the New York City Educational System.)

Major Consulting and Professional Positions (continued)

1964 - 1968	Associate Editor: Public Opinion Quarterly
1964 - 1968	Director: National Science Foundation Research Program in Psychology at Princeton University.

Selected Activities with Professional Societies and Organizations Since 1967

Sept., 1971	Koslin, S., Koslin, B., Pargament, R. & Bird, H. Children's social distance constructs: A developmental study. Paper accepted for presentation to the meetings of the American Psychological Association, Washington, D.C.
Feb., 1971	Koslin, S., Koslin, B. & Pargament, R. Relationships between educational integration policies and students' racial attitudes. Paper presented to the meetings of the American Educational Research Association, New York City.
March, 1970	Koslin, S., Koslin, B., Pargament, R. & Waxman, H. Classroom racial balance and students' interracial attitudes. Paper presented at the meetings of the American Educational Research Association, Minneapolis.
Sept., 1969	Koslin, S., Koslin, B., Cardwell, J. & Pargament, R. A quasi-disguised and structured measure of school-children's racial preferences. Paper presented at the meetings of the American Psychological Association, Washington, D.C.

Selected Activities with Professional Societies and Organizations Since 1967 (continued)

April, 1969	Koslin, B., Cardwell, J. & Pargament, R. Which Negroes prefer what skin-color? Paper presented to the meetings of the Eastern Psychological Association, Philadelphia.
Sept., 1967	Koslin, B. Muzafer Sherif's contributions to social psychology. Invited address: selected to present Kurt Lewin Memorial Award, Society for the Psychological Study of Social Issues, American Psychological Association.
April, 1967	Koslin, B. & Pargament, R. Are assimilation-contrast displacement effects a function of perceptual distortions? Paper presented to the meetings of the Eastern Psychological Association.
April, 1967	Koslin, B. Field research and conceptual systems theory. Discussant's paper, Symposium on Conceptual Theory Research, Eastern Psychological Association.

Invited Colloquia Since 1967

1971	Graduate Division, University of the City of New York.
1970	Columbia University, New York City.
1968	Purdue University.
1968	The University of the City of New York: Brooklyn College.
1967	The University of the State of New York at Albany.

Bertram L. Koslin

Invited Colloquia Since 1967 (continued)

1967	Rutgers, The State University of New Jersey.
1967	Columbia University.
1967	The University of the City of New York.
1967	Rutgers, The State University of New Jersey.

Past and Current Research Grants

Total: 5 (NIH and Peace Corps)

Current: A Study of Race and Attitudes Toward School. Co-Principal Investigators S. Koslin and B. Koslin. Period covered: 1966 - present.

Selected Publications Since 1967

1. Koslin, B.L., Suedfeld, P. & Pargament, R. Belief instability as a moderator variable in opinion change. The British Journal of Social and Clinical Psychology. (In press.)

2. Koslin, S., Koslin, B.L., Pargament, R. & Waxman, H. Classroom racial balance and students' interracial attitudes. Sociology of Education. (In press.)

3. Koslin, B., Pargament, R. & Suedfeld, P. An uncertainty model of attitude change. In P. Suedfeld (Ed.) Alternatives in Attitude Theory. New York: Atherton. In press.

4. Koslin, B.L., & Pargament, R. Effects of attitude on the discrimination of opinion statements. Journal of Experimental Social Psychology, 1969, 5, 254-264.

5. Koslin, B.L. Muzafer Sherif's contributions to social psychology. Journal of Soc. Issues, 1968, 24, 36-37.

Selected Publications Since 1967 (continued)

6. Koslin, B.L., Haarlow, R.N., Karlins, M. & Pargament, R. Predicting group structure from member's cognitions. Sociometry, 1968, 31, 64-75.

7. Koslin, B.L. Laboratory experiments and attitude theory. In C. Sherif & M. Sherif (Eds.) Attitude, ego-involvement and change. New York: Wiley, 1967, pp. 76-87.

8. Sherif, M. & Koslin, B.L. Some comments on the institutional-behavioral controversy with special reference to political science. In M. Sherif (Ed.) Social interaction: Processes and products. Chicago: Aldine, 1967, pp. 98-116.

9. Koslin, B.L., Stoops, J. & Loh, W. Source characteristics and communication discrepancy as determinants of attitude change and conformity. J. Exp. Soc. Psychol., 1967, 3, 230-242.

10. Koslin, B.L., Pargament, R. & Levine, S. Effects of learning on judgment in the presence of discrepant anchors. Psychonomic Science, 1967, 9, 565-566.

Ben Josephson, Jr.

EDUCATION

1959 Fulbright Scholar
University of Nottingham, England

1958 Massachusetts Institute of Technology
Ph.D. Physics

1953 Cornell University
Bachelors of Engineering Physics

PROFESSIONAL EXPERIENCE

1970- RIVERSIDE RESEARCH INSTITUTE
New York City
Acting Director, Social Systems Division

1970 RIVERSIDE RESEARCH INSTITUTE
New York City
Director, Urban Systems Division

1969 - 1970 RIVERSIDE RESEARCH INSTITUTE
New York City
Assistant to the Vice President for Institute Development

1968 - 1969 RIVERSIDE RESEARCH INSTITUTE
New York City
Member of the Research Staff

1969 - 1970 THE TREADWELL CORPORATION
Consultant

1968 - 1969 NEW YORK UNIVERSITY
Adjunct Associate Professor of Physics

1967 - 1968 NEW YORK UNIVERSITY
Associate Professor of Physics

1963 - 1967 NEW YORK UNIVERSITY
Assistant Professor of Physics

1960 - 1963 RICE UNIVERSITY
Houston, Texas
Assistant Professor of Physics

Ben Josephson, Jr.

PROFESSIONAL EXPERIENCE (continued)

1959 - 1960 — RICE UNIVERSITY
Houston, Texas
Research Associate, Low Temperature Group

PROFESSIONAL AND HONORARY SOCIETIES

Tau Beta Pi
Society of the Sigma Xi
American Physical Society
The Institute of Physics and the Physical Society (England)
The Institute of Electrical and Electronic Engineers
The New York Academy of Sciences
The American Association for the Advancement of Science
Listed in American Men of Science

PUBLICATIONS

Open Literature:

The coupling of a spin system to a cavity mode. Proc. Phys. Soc., 1959, 74, 561. (with K.W.H. Stevens)

The effect of color centers on the nuclear spin-lattice relaxation time of F nuclei in LiF. J. Phys. Chem. Solids, 1962, 23, 67. (with M.W.P. Strandberg)

Microwave ultrasonic attenuation in ruby. Bull. Am. Phys. Soc., 1962, 1, 15. (with R.N. Claytor and P.L. Donoho)

Nuclear spin-lattice relaxation time in CaF_2 doped with rare earth elements. Bull. Am. Phys. Soc., 1963, 8, 346. (with M.A. Waldrup)

Phonon-phonon interaction in CaF_2. Bull. Am. Phys. Soc., 1964, 9, 533. (with B.R. Breed and P.L. Donoho)

Spin-lattice relaxation of F^{19} in CaF_2 in low temperature. Phys. Rev., 1965, 137, A108. (with S.M. Day and E. Otsuka)

PUBLICATIONS (continued)

Numerous RRI reports and proposals, for example:

Study of the Emergency Reporting System for New York City.

Cable Television in the Public Interest: New Directions for New York State.

A Targeting System for Occupational Education.

Samuel A. Scharff

EDUCATION

1943 Massachusetts Institute of Technology
B.S.E.E.

1964 Stevens Institute of Technology
MS
Interdepartmental program of courses relevant to design and use of digital computers.

Massachusetts Institute of Technology
SBEE, and two terms of Graduate School courses.
Program centered on feedback control systems.

SUMMARY WORK EXPERIENCES

1970 - present — RIVERSIDE RESEARCH INSTITUTE
New York City
Assistant to the Vice President of Research.

August, 1968 - 1970 — RIVERSIDE RESEARCH INSTITUTE
New York City
Systems Engineer. Programs analyst, programming, planning and budgeting. Management monitor of research

January, 1953 - August, 1968

"Data/Control Systems": computers, automatic controls, communications, displays, and control panels -- integrated systems for men with a job to do.

Requirements studies; feasibility studies; designs, including cost/effectiveness comparisons; construction; test; maintenance.

Also, components for Data/Control Systems: mechanical, optical, electrical, and electronic devices -- design, construction, test, maintenence.

Samuel A. Scharff

SUMMARY WORK EXPERIENCES (continued)

Clients:

President's Task Force on Communications Policy
United States Air Force - Office of the Assistant Secretary, R & D; Rome Air Development Center, A. F. Systems Command
The RAND Corporation
General Electric Company
The Louis Allis Company
Automation Dynamics Corporation
Capehard Corporation
Dennison Manufacturing Co.
Jamaica Water Supply Co.
Royal McBee Corporation
University Controls Corporation
James P. O'Donnell, Consulting Engineer

STEVENS INSTITUTE OF TECHNOLOGY
Visiting Lecturer

PREVIOUS WORK EXPERIENCE

1941 - 1953

M. W. Kellogg Company.
Project Engineer, Military Service, U.S. Air Force: Rome Air Development Center; Far East Air Forces.
General Electric Co., 12 months of full-time work while in M.I.T. cooperative course plan in Electrical Engineering.

PROFESSIONAL AFFILIATIONS

Institute of Electrical and Electronics Engineers
Association for Computing Machinery
Society for General Systems Research
New York State and National Societies of Professional Engineers

LICENSED PROFESSIONAL ENGINEER: State of New York and State of New Jersey.

Richard Pargament

EDUCATION

1965 Bard College
A.B.

1967 Princeton University
M.A. Psychology

1968 Princeton University
Ph.D. Psychology

ACADEMIC AND PROFESSIONAL EXPERIENCE

1969 - Present

RIVERSIDE RESEARCH INSTITUTE
New York City
Manager, Social Systems Division

1968 - 1969

RUTGERS - THE STATE UNIVERSITY
Assistant Professor; Department of Psychology

RIVERSIDE RESEARCH INSTITUTE
New York City
Consultant

1965 - 1968

PRINCETON UNIVERSITY
Teaching Fellow: Statistics, Personality, Social Psychology

Research Assistant, Communication and Social Interaction Laboratory

RUTGERS - THE STATE UNIVERSITY
Coadjutant Lecturer, Department of Psychology

INSTITUTE OF EDUCATIONAL DEVELOPMENT
Psychometric Consultant

PROFESSIONAL MEMBERSHIPS

American Educational Research Association
American Psychological Association

Richard Pargament

PUBLICATIONS

Belief instability as a moderator variable in opinion change. The British Journal of Social and Clinical Psychology, in press. (with B.L. Koslin and P. Suedfeld)

An experimental attack on smoking. International Journal of the Addictions, in press. (with P. Suedfeld, P.B. Landon, and Y.M. Epstein)

Classroom racial balance and students' interracial attitudes. Sociology of Education, in press. (with S. Koslin, B.L. Koslin and H. Waxman)

An uncertainty model of attitude change. In P. Suedfeld (Ed.) Alternatives in Attitude Theory. New York: Atherton, 1972. (with B.L. Koslin and P. Suedfeld)

The role of experimenter and subject expectations in sensory deprivation. Representative Research in Social Psychology, 1971, 2 no. 1. (with P. Suedfeld, P.B. Landon and Y.M. Epstein)

Effects of attitude on the discrimination of opinion statements. Journal of Experimental Social Psychology, 1969, 5, 245-264. (with B.L. Koslin)

Predicting group structure from member's cognitions. Sociometry, 1968, 31, 64-75. (with B.L. Koslin, R.N. Haarlow and M. Karlins)

Effects of learning on judgment in the presence of discrepant anchors. Psychonomic Science, 1967, 9, 565-566. (with B.L. Koslin and S. Levine)

RECENTLY PRESENTED PAPERS

Efficacy of school integration policies in reducing racial polarization. American Psychological Association symposium paper, to be presented September, 1972. (with S. Koslin and B.L. Koslin)

Children's social distance constructs: a developmental study. American Psychological Association paper, Washington, Sept., 1971. (with S. Koslin, B.L. Koslin and H. Bird)

RECENTLY PRESENTED PAPERS (continued)

Relationships between educational integration policies and students' racial attitudes. American Educational Research Association paper, New York City, February, 1971. (with S. Koslin and B.L. Koslin)

Classroom racial balance and students' interracial attitudes. American Educational Research Association paper, Minneapolis, Minn., 1970. (with S. Koslin, B. Koslin and H. Waxman)

Which Negroes prefer what skin color? Eastern Psychological Association paper, Philadelphia, 1969. (with B.L. Koslin and J. Cardwell)

A quasi-disguised and structured measure of schoolchildren's racial preferences. American Psychological Association paper, Washington, 1969. (with B.L. Koslin, S.C. Koslin and J. Cardwell)

Are assimilation contrast displacement effects a function of perceptual distortions? Eastern Psychological Association paper, April, 1967. (with B.L. Koslin)

Henry B. Bird

EDUCATION

1963 Bard College
A.B.
1967 University of Oklahoma
M.A. Sociology
1968 University of Oklahoma
Ph.D. Social Psychology

PROFESSIONAL EXPERIENCE

1972 - Present — RIVERSIDE RESEARCH INSTITUTE
New York City
Assistant Manager, Social Systems Division

1970 - 1972 — RIVERSIDE RESEARCH INSTITUTE
New York City
Member of the Research Staff

1968 - 1969 — CENTER FOR COMMUNITY RESEARCH
New York City
Director, Computer Applications Department

1965 - 1968 — THE UNIVERSITY OF OKLAHOMA MEDICAL CENTER
Norman, Oklahoma
Research Assistant for Dr. John Bruhn

1963 - 1965 — INSTITUTE OF GROUP RELATIONS
University of Oklahoma
Research Fellow for Dr. Muzafer Sherif

AREAS OF EXPERIENCE

Research design and evaluation of projects in social welfare, medical sociology, learning processes, small group behavior, attitudes and attitude change, and public education.

Interviewing and participant observation experience.

Henry B. Bird

AREAS OF EXPERIENCE (cont'd)

Experience in Fortran, PL1, Cobol, OS, DOS, 360-series computers, data base management, and optical scanning technology.

PUBLICATIONS

Social aspects of coronary heart disease in a Pennsylvania German community, Social Science and Medicine, 1968, 2, 201-212. (with J. Bruhn and B. Chandler).

Some selected factors relating to adolescent group functioning, published Doctoral dissertation, Univ. of Oklahoma, August, 1968.

Social profiles and academic standing: A study of first year medical students. The Oklahoma State Medical Journal, October, 1967. (with Bruhn and Adsett)

A study of the class distribution of small solid adolescent groups, unpublished Master's thesis, University of Oklahoma, 1966.

A preliminary study of inter-clique relations, Bard Psychology Journal, 102, August, 1963.

PRESENTED PAPERS

Children's social distance constructs: A developmental study. American Psychological Association paper, Washington, Sept., 1971. (with S. Koslin, B.L. Koslin and R. Pargament)

PROFESSIONAL MEMBERSHIP

American Sociological Association

Sandra Cohen Koslin

EDUCATION

1959 Queens College, City University of New York
A.B. Elementary Educ.

1961 Teachers College, Columbia University
A.M. Educational Psych.

1964 Teachers College
Columbia University
Ph.D. Educational Psych.

ACADEMIC AND PROFESSIONAL EXPERIENCE

1968 - Present
RIVERSIDE RESEARCH INSTITUTE
New York City
Member of the Research Staff

1968 - 1969
INSTITUTE FOR COMMUNITY STUDIES
Sarah Lawrence College
Bronxville, New York
Research Psychologist and Teaching Associate

1964 - 1968
EDUCATIONAL TESTING SERVICE
Princeton, New Jersey
Research Psychologist

1963 - 1964
LEXINGTON SCHOOL FOR THE DEAF
New York City
Research Assistant

1961 - 1963
TEACHERS COLLEGE, COLUMBIA UNIVERSITY
New York City
Research Assistant

1959 - 1960
NEW YORK CITY PUBLIC SCHOOLS
East Harlem in New York City
4th grade Teacher

Sandra Cohen Koslin

HONORS

New York Regents Scholar
Phi Beta Kappa
Summa Cum Laude
National Defense Fellow
Psi Chi
Sigma Xi

RECENT PUBLICATIONS

1. Koslin, S., Koslin, B.L., Pargament, R. & Waxman, H. Classroom racial balance and students' interracial attitudes. Sociology of Education, in press, 1972.
2. Koslin, S.C., Amarel, M. & Ames, N. The effect of race on peer evaluation and preference in primary grade children: An exploratory study. Journal of Negro Education, 1970, 39, 346-350.
3. Koslin, S.C., Amarel, M. & Ames, N. A method for assessing primary grade children's expectations of school. ETS Research Bulletin 70-25. Princeton: Educational Testing Service, 1970.
4. Koslin, S.C. Equal educational opportunity. Education and Urban Society, 1969, 2, 119-121.
5. Koslin, S.C., Amarel, M. & Ames, N. A distance measure of racial attitudes in primary grade children: An exploratory study. Psychology in the Schools, 1969, 6, 382-385.
6. Cohen, S.R. An exploratory study of young children's attitudes toward school. Proceedings, 75th Annual APA Convention, 1967.
7. Cohen, S.R. Exploratory study of young children's perceptions of teacher behavior. Proceedings, 75th Annual APA Convention, 1967.
8. Cohen, S.R. Predictability of deaf and hearing story paraphrases. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 916-921.
9. Cohen, S.R. An exploratory study of student attitudes in the primary grades. ETS Research Bulletin 65-30. Princeton: Educational Testing Service, 1965.
10. Cohen, S.R. Redundancy in the written language of the deaf: Predictability of story paraphrases written by deaf and hearing children. In Rosenstein, J. and MacGinitie, W. (Eds.) Research studies on the psycholinguistic behavior of deaf children. Washington: The Council for Exceptional Children, NEA, 1965.

Sandra Cohen Koslin

RECENT PAPERS PRESENTED AT LEARNED SOCIETIES

1, 2. Koslin, S., Koslin, B.L. & Pargament, R. Efficacy of school integration policies in reducing racial polarization. American Psychological Association, Washington, D.C., September, 1971 and 1972.

3. Koslin, S., Koslin, B.L., Pargament, R. & Bird, H. Children's social distance constructs: A developmental study. American Psychological Association, Washington, D.C., September, 1971.

4. Koslin, S., Koslin, B.L. & Pargament, R. Relationships between educational integration policies and students' racial attitudes. American Educational Research Association, New York City, February, 1971.

5. Koslin, S., Koslin, B.L., Pargament, R. & Waxman, H. Classroom racial balance and students' interracial attitudes. American Educational Research Association, Minneapolis, March, 1970.

6. Koslin, S., Koslin, B.L., Cardwell, J. & Pargament, R. A quasi-disguised and structured measure of schoolchildren's racial preferences. American Psychological Association, Washington, D.C., September, 1969.

7. Koslin, S.C. The measurement of schoolchildren's racial attitudes: A validity study. Eastern Psychological Association, Philadelphia, April, 1969.

8. Koslin, S., Amarel, M. & Ames, N. A distance measure of social perceptions in white and Negro primary grade children. American Psychological Association, San Francisco, September 1968.

9. Koslin, S., Amarel, M. & Ames, N. The effect of race on peer evaluation and preference in primary grade children: An exploratory study. American Educational Research Association, Chicago, February, 1968.

Jerry P. Brashear

EDUCATION

1967 Princeton University
Princeton, New Jersey
B.A. Psychology

1969 Harvard Business School
Boston, Massachusetts
M.B.A.

Present University of Michigan
Ann Arbor, Michigan
Ph.D. Candidate - Urban and Regional Planning

ACADEMIC AND PROFESSIONAL EXPERIENCE

1971 - present — RIVERSIDE RESEARCH INSTITUTE
New York City, New York
Member of the Research Staff

1969 - 1971 — COMMUNITY SYSTEMS FOUNDATION (CSF)
Ann Arbor, Michigan
Consultant

summer 1969 — MC KINSEY AND COMPANY, INC.
New York City, New York
Research Associate

summer 1968 — MAYOR'S OFFICE, CITY OF NEW YORK
New York City, New York
Mayor's Intern

AREAS OF EXPERIENCE

Program planning of the Community Renewal Program for the City of Ann Arbor. Worked on defining problems in physical planning, capital budgeting, and PPBS; designing research and analytic steps; design and implementation of project management/control system.

Analysis of program plans for Harambee, Inc., a black community development agency: cashflow planning, organizational structure, implementation tactics.

Planning, management, and analytic consultation on various departmental problems within selected local government agencies.

Analysis of New York City government organizational structure: Mayor's Office and staff-line relations, community control questions addressed.

AREAS OF EXPERIENCE (cont.)

Analysis of the budgetary and control processes of the Boston Police Department: budget preparation and justification; cost accounting, management planning and control; analytic capabilities.

Summary investigation of the present and projected involvement of the private sector in public education; street academies and training programs; computer - aided instruction; curriculum development, etc.

Experimental research into the affects on attitude change of source prestige, discrepancy, and reference scale stability, singly and in interactions.

CORNELIUS LANGLEY

Education

1949 Columbia University
B.A. Mathematics
1951 Columbia University
M.A. Mathematics

Professional Experience

August 1972 - present — RIVERSIDE RESEARCH INSTITUTE
New York City
Member of the Research Staff

August 1967 - August 1972
(6/65 - 9/66 part-time) — MOBIL OIL CORPORATION
New York City
Associate Engineer and
Consultant in Mathematics

September 1966 - August 1967 — NATIONAL SCIENCE FOUNDATION
Fellowship in Advanced Mathematics

September 1952 - August 1967 — NEW YORK UNIVERSITY
New York City
Assistant Professor of Mathematics

February 1949 - September 1952 — COLUMBIA UNIVERSITY
New York City
Lecturer in Mathematics

Honorary Societies

Phi Beta Kappa

Philip Bruce Landon

EDUCATION

1967 University of Utah, B.S.

1969 Rutgers University
M.S. Psychology

1972 Rutgers University
Ph.D. Psychology

ACADEMIC AND PROFESSIONAL EXPERIENCE

1972 - present	RIVERSIDE RESEARCH INSTITUTE New York City Member of the Research Staff
1969 - 1972	BELL TELEPHONE LABORATORIES Systems Analyst
1969 - 1972	UNIVERSITY COLLEGE, RUTGERS - THE STATE UNIVERSITY Co-adjutant Instructor
1969 - 1970	RUTGERS - THE STATE UNIVERSITY N.I.M.H. Predoctoral Fellow
1968 - 1969	RUTGERS - THE STATE UNIVERSITY Teaching Asst. for University College Psychology Dept.
1968	RUTGERS - THE STATE UNIVERSITY Research Assistant
1967	CITY OF NEW BRUNSWICK, NEW JERSEY Research Assistant, (Urban Renewal Project)
1967	UNIVERSITY OF UTAH N.S.F. Undergraduate Research Fellow
1967	PEACE CORPS, IRAN TRAINING PROJECT Research Assistant, Peace Corps

Philip Bruce Landon

ACADEMIC AND PROFESSIONAL EXPERIENCE (continued)

1967	CLEARFIELD JOB CORPS INSTITUTE Clearfield, Utah Research Assistant in measurement of human relations skills
1967	UNIVERSITY OF UTAH Research Assistant for Esso Creativity in Engineering Project
1966	UNIVERSITY OF UTAH Research Assistant

PROFESSIONAL MEMBERSHIPS

American Psychological Association

PUBLICATIONS

Complex cognitive performance and sensory deprivation: Completing the U-curve. Perceptual and Motor Skills, in press. (with P. Suedfeld)

An experimental attack on smoking. (Attitude manipulation in restricted environments, III.) International Journal of the Addictions, in press. (with P. Suedfeld, R. Pargament, and Y.M. Epstein)

The role of experimenter and subject expectations in sensory deprivation. Representative Research in Social Psychology, 1971, 2 21-27. (with P. Suedfeld, Y.M. Epstein and R. Pargament)

The effects of set on mere exposure. Journal of Personality and Social Psychology, 1971, 17 121-123. (with P. Suedfeld, Y.M. Epstein and P. Buchanan)

Motivational arousal and task complexity: support for a model of cognitive changes in sensory deprivation. Journal of Experimental Psychology, 1970 83 329-330. (with P. Suedfeld)

Information and meaningfulness needs in sensory deprivation. Psychonomic Science, 1969, 5 (4) 248. (with P. Suedfeld)

Lucille M. McCulley

EDUCATION

1950 Duchesne College
A.B. English & French

1951 Faculté de Lettres
(Rennes, France)
Certificate for License
in Modern History

1952 Institut d'Etudes Politiques, Univ. of Paris

1953 Columbia University
M.A. Public Law and
Government

1956 New York Institute of
Finance

AREAS OF EXPERIENCE

1972 - Present — RIVERSIDE RESEARCH INSTITUTE
New York City
Member of the Research Staff

1958 - 1972 — OPERATIONS MANAGER-ELECTRONICS
New York City
Executive Vice President

1954 - 1958 — WHITE, WELD & CO.
New York City
Administrative Assistant
and Registered Representative

1952 - 1954 — COLUMBIA UNIVERSITY
New York City
Research Assistant in
Development Office

HONORS

Fulbright Scholarship, France (1950-52)